WINE PRICES AND QUALITY RATINGS: A META-REGRESSION ANALYSIS

EDWARD OCZKOWSKI AND HRISTOS DOUCOULIAGOS

This article examines the empirical support for the hypothesized hedonic theoretical relation between the price of wine and its quality. The examination considers over 180 hedonic wine price models developed over 20 years, covering many countries. The research identifies that the relation between the price of wine and its sensory quality rating is a moderate partial correlation of +0.30. This correlation exists despite the lack of information held by consumers about a wine's quality and the inconsistency of expert tasters when evaluating wines. The results identify a moderate price-quality correlation, which suggests the existence of strategic buying opportunities for better informed consumers. Strategic price setting possibilities may also exist for wine producers given the incomplete quality information held by consumers. The results from the meta-regression analysis point to the absence of any publication bias, and attribute the observed asymmetry in estimates to study heterogeneity. The analysis suggests the observed heterogeneity is explained by the importance of a wine's reputation, the use of the 100-point quality rating scale, the analysis of a single wine variety/style, and the employed functional form. The most important implication from the analysis is the relative importance of a wine's reputation over its sensory quality, inferring that producers need to sustain the sensory quality of a wine over time to extract appropriate returns. The reputation of the wine producer is found not to influence the strength of the price-quality relationship. This finding does not contradict the importance of wine producer reputation in directly influencing prices.

Key words: Hedonic wine prices, incomplete information, meta-regression, wine quality.

JEL codes: C21, Q11.

Examining the relationship between the price of a product and its quality has a long history that continues to command significant interest in the economics literature (e.g., McConnell and Strand 2000; Teuber and Herrman 2012). One of the earliest attempts at measuring the relationship dates to Waugh's (1928) work on the relation between the price of vegetables and their quality, while the foundation of economic theory associated with establishing the price-quality relationship can be traced to Rosen's (1974) hedonic model. The model views the demand and supply for product attributes as interacting to determine market prices. The consequent hedonic price function is a relationship between the price of a product and its attributes.

Wine lends itself well to hedonic price analysis because of the substantial attribute variation in available wine products. Wine products that employ grapes from different varieties, grown in different regions, and at different time points are typically available for purchase. Even for similar products (i.e., those using grapes from the same variety, region, and vintage), wines can differ significantly because of grape quality variations and the practices of different wine producers. Oczkowski (1994) provided one of the first hedonic price studies for wine when he examined the relation between price and the quality ratings from a wine guide, and a series of other "objective" attributes such as the region, variety, and vintage of the employed grapes for Australian premium wines.

The number of hedonic wine price studies has subsequently grown, with over sixty

Edward Oczkowski is a professor at the School of Accounting and Finance, Charles Sturt University, Australia. Hristos Doucouliagos is a professor at the School of Accounting, Economics, and Finance, Deakin University, Australia. The authors acknowledge the assistance of Guenter Schamel, Wei Zhao, and Karl Storchmann for providing information about some standard errors of published estimates. The comments from two referees and the editor Madhu Khanna are gratefully acknowledged. Correspondence may be sent to: eoczkowski@csu.edu.au.

Amer. J. Agr. Econ. 97(1): 103–121; doi: 10.1093/ajae/aau057
Published online July 13, 2014

identifiable studies (e.g., Combris, Lecocq, and Visser 1997; Oczkowski 2001; and Benfratello, Piacenza, and Sacchetto 2009). These studies cover wines from all over the world using varying sample designs and modifying the basic model in a whole series of directions. Extensions of the basic framework have examined the role of individual sensory quality variables, chemical measurements, reputation (wine, producer, and region), vineyard climatic variables, measurement error, the endogeneity of quality ratings, segmentation of price ranges, investment ratings, and cases produced. For a summary of various estimated models, see Fogarty (2003) and Estrella Orrego, Defrancesco, and Gennari (2012).

Examining the price and quality relationship is important in the wider product differentiation context in mature markets. Significant empirical literature exists examining the market characteristics that impact on the price and quality relationship (e.g., Caves and Greene 1996; and Volckner and Hofmann 2007). Moreover, theoretical models have been developed that identify appropriate price and quality-setting strategies by firms in the presence of incomplete information (e.g., Daughety and Reinganum 2008; and Schnabel and Storchmann 2010).

An important debate in the wine price and quality literature surrounds the following question: Does quality matter? The nature of the debate stresses the imprecise and subjective nature of wine quality assessment, and particularly consumers' perception of wine quality. In summarizing the price-quality debate, Lecocq and Visser (2006) suggest that "it has always been animated and controversial." The most-cited paper in the literature (Combris, Lecocq, and Visser 1997) finds that quality does not matter when prices are regressed against individual sensory variables and other objective wine characteristics. In contrast, a significant number of other studies (e.g., Landon and Smith 1998; and Schamel 2009) have identified the importance of a wine guide's overall sensory rating in determining price variations.

The purpose of the study is to address this fundamental research question of whether there exists a strong empirical relationship between a wine's price and its quality. The analysis also addresses a related sub-question: What research design factors account for the observed heterogeneity in the reported price-quality relationship? To answer these questions, meta-regression analysis (MRA; see Stanley and Doucouliagos 2012), will be employed to estimate the effect size of the price-quality relationship across a series of studies. Meta-regression analysis also allows us to model the inherent heterogeneity across studies. This allows us to control for the effects of study design and use different extraneous variables when estimating the relationship between price and quality. We also modify the standard MRA model by using atypical standard errors that correct for both data dependence and the unequal number of estimates reported by individual studies. This correction is rare in MRA, and our results show that correction can make a difference.

In the next section we discuss the general theoretical foundation for the relation between price and wine quality. We focus on a series of key issues such as the hedonic foundation, quality measurement, and information flow. Subsequently, the data for the meta-analysis is outlined and an initial examination of the relation between price and quality is undertaken. We then present and discuss the MRA to identify how this relation is impacted by various sample design and extraneous variable factors.

The Wine Price and Quality Relationship

Most price-quality wine studies cite Rosen's (1974) hedonic framework as their theoretical foundation. Rosen assumes that a product consists of a series of attributes that demanders desire and suppliers are willing to produce. For wine, Verdu Jover et al. (2004), Charters and Pettigrew (2007), and others suggest that wine quality is a multidimensional construct comprised of both extrinsic and intrinsic attributes. The extrinsic factors include the grape type, reputation, fault-free wine production, and marketing aspects. The intrinsic factors are aroma, appearance, and gustatory factors. The intrinsic factors may be measurable through the sensory evaluation of traits such as complexity, balance, personality, length, varietal purity, and intensity of flavor. Even though wine is desired for a whole series of attributes, our focus rests on its overall sensory quality.

The Rosen (1974) framework suggests that demanders express a bid function for the sensory quality attribute. This represents the maximum price they are willing to pay for a level of sensory quality. The bid function is assumed to be increasing in sensory quality at a decreasing rate, that is, demanders are willing to pay more for increasing levels of sensory quality. Symmetrically, suppliers express an offer function for the sensory quality attribute. This represents the minimum price they are willing to accept for a sensory quality level. The offer function is assumed to be increasing in sensory quality at a decreasing rate, that is, suppliers are willing to accept a lower price only for decreasing sensory quality levels. Assuming perfect competition (including perfect information flows), in equilibrium the various bid and offer functions are tangential. These tangents trace out the market hedonic price function with reference to the sensory quality attribute. The model predicts that higher sensory quality levels are associated with higher prices.

To motivate the importance of sensory quality in the bid and offer functions, we consider the literature in oenology, viticulture, sensory psychology, and wine marketing. This literature attempts to identify the concept of wine sensory quality through taste preferences. The literature also describes how wine quality relates to vineyard practices and wine making procedures. For example, Varela and Gambaro (2006) and Saenz-Novajas et al. (2010) provide some evidence to establish the existence of the relationship between sensory traits and overall wine quality ratings made by expert wine tasters. This relation between quality ratings and sensory traits has also been established in the economics literature by Combris, Lecocq, and Visser (1997) and Lecocq and Visser (2006) through the estimation of "jury equations." These studies provide some justification for consumer bid functions that depend on sensory quality; consumers prefer higher levels of desirable sensory traits and this results in higher price bids.

For grape and wine producers, research has examined the relation between wine sensory quality and vineyard and wine making practices (Reynolds 2010). For example, Smart (1985) and Jackson and Lombard (1993) discuss the importance of canopy management, vine yield, leaf removal, and vine thinning in producing better quality grapes. Diaz-Plaza et al. (2002) and Garde-Cerdan and Ancin-Azpilicueta (2006) analyze the impact of oak in maturation on wine quality. These studies provide some justification for the existence of supplier offer functions that depend on sensory quality. These vineyard and wine making practices that increase quality are more costly to implement and thus result in higher price offers.

It is important to recognize that the hedonic price function is justified through the interaction of both consumers and producers, and not consumers alone. A focus purely on wine consumers should only result in the estimation of bid functions, which may not be identifiable from market transaction data. The taste preferences of consumers and the difficulty (costs) in producing these desirable quality attributes both play a role in establishing the market relationship between the price and the sensory quality of wine.

A key assumption of the hedonic framework is that both demanders and suppliers have perfect knowledge about the attributes or quality of the product, and that these characteristics are "objectively measureable" (Rosen 1974). Given the subjective nature of wine quality, the framework possibly needs modification to be applied to wine products. The expertise of wine makers and their use of chemical analysis (measuring sugar and acid levels, etc.) and sensory evaluations in producing wine implies that suppliers have a good understanding of the quality of the wine product they offer for sale. Many studies (e.g., Gawel, Royal, and Leske 2002; Parr, White, and Heatherbell 2004; and Ballester et al. 2008) suggest that wine experts (including winemakers) are better at evaluating the quality attributes of wine than novice consumers. This evidence provides some justification for the use of a sensory quality measure of the wine appearing in the hedonic price function.

In contrast, there is a less certain understanding of wine quality held by wine consumers.[1] Many authors (e.g., Ali and Nauges 2007) argue that wine is an experience good whose quality and therefore desirability can only be assessed after consumption. Further, a consumer must form an expectation of the quality of the wine before purchase. This recognition has led to at least two types of practices for estimating hedonic price functions.

First, in the price function, rather than employing a sensory quality measure of the

[1] This recognition of lacking consumer information about wine quality is used to explain the statistical unimportance of individual sensory variables in hedonic price functions (see Combris, Lecocq, and Visser 1997: Lecocq and Visser 2006).

wine, the reputation of the wine (possibly based on past measures of sensory quality) is used (e.g., Landon and Smith 1997, 1998; Oczkowski 2001; and Zhao 2008). This argument also extends to motivate the inclusion of producer and regional reputation variables in the price function (see Landon and Smith 1998; Schamel 2009). An important observation however, is that this argument only questions the use of a sensory quality measure in the hedonic price function from the demand side of the market. The motivation for including a quality variable remains from the supply side.

Second, some measure of sensory quality is employed and sourced from a publicly accessible wine guide. Some researchers argue that reviews of wines act as opinion leadership (Edwards and Mort 1991), and may influence wine consuming practices. Johnson and Bruwer (2004) argue that the use of wine guides by consumers acts as a risk reduction strategy in the absence of perfect information. These authors show that wine reviews do drive a significant portion of consumer behavior. Further, Friberg and Gronqvist (2012) illustrate how significant demand increases for wines do occur after the release of favorable reviews.

Several counter-arguments against the importance of wine guides in influencing demand and hence prices can be articulated. First, a series of studies have showed a disconnect between the taste preferences of wine experts and novice wine consumers (Schiefer and Fischer 2008; Lattey, Bramley, and Francis 2010; and D'Alessandro and Pecotich 2013). Second, as argued by Landon and Smith (1998), Benfratello, Piacenza, and Sacchetto (2009), and others, wine guides are typically released with information on prices and so a quality rating from a guide may be determined after setting the recommended retail price (RRP). In these circumstances, the RRP cannot be influenced by publicly available quality ratings. Though this argument has some merit, it may be mitigated by the recognition that some wine producers may have some prior warning of reviews and therefore form recommended prices in anticipation of demand.[2] Further, some researchers do employ an average actual retail price (formed sometime after wine guide release) that will capture any real price adjustments in response to excessive or weak demand. In other cases, researchers may be able to employ auction prices. Auction prices will better reflect demand influences from wine guides over time for more exclusive wines.

Another potential drawback with the use of sensory expert quality ratings in developing the price-quality relationship is the inconsistency between various expert judges (consensus) and the inconsistency of the same expert judge over time (reliability) when assessing sensory quality. When summarizing the literature on the consistency of experienced wine judges, Ashton (2012) finds, for various studies (e.g., Hodgson 2008), that the mean reliability between judges is 0.50. The mean consensus over time for the same judge is estimated to be 0.34.[3] These measures of consistency are found to be lower than those used in other disciplines and therefore may have implications for the stability of the price-quality relationship across studies that employ different expert ratings. The potential impact of the lack of consistent expert ratings also focuses some attention on the type of quality rating scales and sample design used to develop the price-quality relationship.

Finally, some literature examines the relationship between prices and consumers' taste preferences, and reveals some conflicting evidence. On the one hand, Goldstein et al. (2008) demonstrate a disconnect between the price of some wines and the taste preference of consumers. In contrast, Mastrobuoni, Peracchi, and Tetenov (2013) provide evidence to suggest that for some consumers prices do signal quality. This conflicting evidence again strengthens the need to examine, across a series of studies, the relationship between price and quality, and to assess the robustness of any such correlation.

In summary, the hedonic framework provides a strong motivation for the wine price-sensory quality relationship. The employed wine sensory quality measure should reflect consumers' willingness to pay more for higher quality wines and suppliers' greater difficulty (costs) in producing higher quality wines. Despite the difficulty in objectively and consistently assessing the sensory quality

[2] An example is provided by Australia's Penfolds Grange 2008. Prior to release it had its recommended retail price raised by 25% in anticipation of high demand after the release of a 100-point review in the *Wine Advocate* (Weekly Times 2013).

[3] These measures employ the average of the correlations between expert scores.

of wine, using wine ratings from publicly accessible expert wine guides may capture these demand and supply influences on wine prices. In part, the purpose of this meta-analysis is to empirically assess these claims and measure the strength of the price-quality relationship.

Meta-regression Analysis Framework and Data

The focus of the meta-regression analysis is the relation between the price and quality of a bottle of wine as measured by a single rating based on a sensory evaluation. This focus rules out a small number of hedonic wine price studies that include other factors but exclude a measure of overall sensory wine quality. Some of the excluded price function estimates focus on wine reputation (e.g., Wade 1999), vineyard climatic conditions (e.g., Ashenfelter 2008), the price of unfinished wines in barrels (e.g., Ali and Nauges 2007), or employ a series of individual sensory and/or chemical attributes (e.g., Combris, Leccoq, and Visser 1997).

The Meta-data

We followed the Meta-Analysis of Economics Research Network (MAER-NET) guidelines in composing the database and conducting the MRA (see Stanley et al. 2013).[4] We searched the following databases using combinations of the keywords (wine, prices, hedonic, and quality) to identify relevant studies: EBSCOhost (which includes Academic Search Complete, Business Source Complete, and Econlit); Emerald Management Complete; Google Scholar; Science Direct; Scopus; Web of Science; and Wiley Online Library. The database searches were conducted during March and April 2013.

In total, we identified 43 separate published studies that estimated the relation between price and an overall wine sensory rating. In MRA, various effect size measures are used including correlations, elasticities, and marginal impacts. We chose the partial correlation coefficient (r) as the effect measure. The partial correlation is an often-proposed measure for meta-analysis (Doucouliagos 1995; Djankov and Murrell 2002) that only needs the t-ratio and the degrees of freedom associated with the estimate for evaluation. As a result, many more estimates are available for analysis than if using an elasticity or other measure. Further, the partial correlation has a common and natural interpretation across studies (Stanley and Doucouliagos 2012).

We did not employ the marginal impact as the effect measure because the employed studies used a variety of non-comparable quality-rating scales (see Cicchetti and Cicchetti 2009) and different functional forms. This made comparisons of marginal impact estimates challenging. We did not employ the elasticity for similar reasons. Further, given the extensive use of the log-linear form, the majority of elasticities can only be evaluated at the means of the data, if such data is available. This lessens the quality and number of available estimates.

Six of the identified studies failed to provide sufficient information to determine the standard error (*se*) of the appropriate effect estimates, which is needed for a valid meta-analysis.[5] Of the remaining studies, an inspection of the data and funnel plots identified two outliers.[6] These were removed for the subsequent analysis of 184 partial correlations from 36 studies. We confirmed that the estimates included in the meta-dataset are comparable and representative.[7] We employ all estimates in the meta-regression analysis

[4] The guidelines for MRA relate to a series of issues including research questions, effect size, literature searching, coding, and modelling. Information on all of these aspects are provided in the article. In terms of data coding, the first author did the initial data coding, with constant double checks. The first author has significantly contributed to the literature by being one of the most-cited authors in the area, and hence has the appropriate expertise to develop the dataset. The second author did some random re-coding to determine the accuracy of the data set.

[5] We contacted authors directly for studies that failed to provide standard errors for estimates. This improved the coverage of studies. Some studies were excluded because they used dummy variables to represent the quality variable for estimation. The arbitrariness of leaving out a control group for estimation makes the calculation of meaningful standard errors problematic.

[6] One outlier was found to be a typographical error. The other outlier was an estimate from a grossly under-specified simple price-quality relation based on a very large data set (covering numerous wine types and regions) that failed to include any appropriate control variables.

[7] We confirmed data comparability by running two tests. First, we tested whether the reported estimates differ according to the quality of the journal in which they were reported. We used the 2012 Social Science Citation Index (SSCI) Journal Impact Factors as proxies for study quality, assigning a zero weight to any journal not indexed in the SSCI. Second, we regressed the precision of the estimated partial correlations against the same Impact Factors. We found no significant difference in partial correlations on the basis of the quality of the journal as measured by Journal Impact Factors.

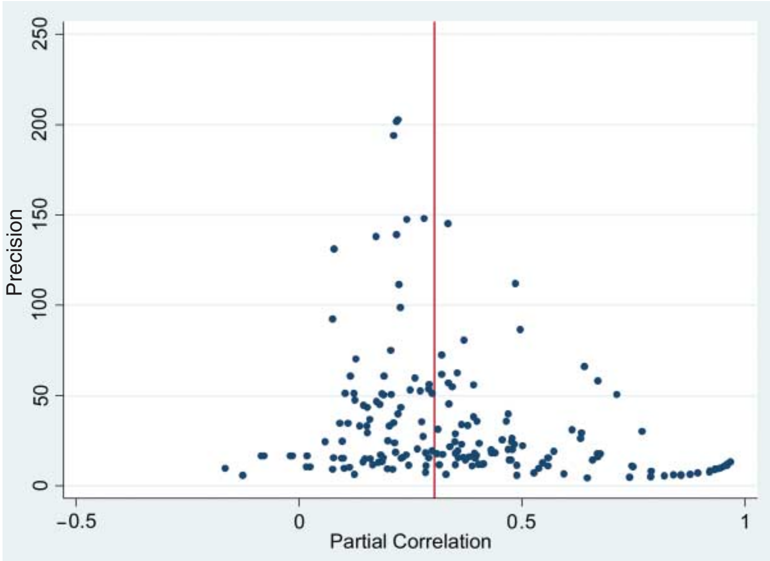

Figure 1. Funnel plot for quality ratings and wine prices partial correlations

Note: The solid line denotes the position of precision weighted average partial correlation (+0.30). The vertical axis measures precision calculated as the inverse of the standard error of the partial correlation.

and explicitly recognize the multiple estimates from some studies by using suitable panel data and cluster error estimation techniques. The details of the studies employed to produce the data and summaries of the effect sizes from each of the studies are provided in the supplementary on-line appendix.[8]

Figure 1 presents the funnel plot of the distribution of reported partial correlations and illustrates the association between the estimated effect size and its precision. Precision is measured as the inverse of the associated standard error (Stanley and Doucouliagos 2010; 2012).[9] Except for a small number of estimates (six), all the estimated partial correlations are positive, showing positive correlations between quality ratings and price. The weighted average of all estimates is +0.30.[10] According to Cohen's (1988) guidelines for the practical significance of a *simple* correlation, a correlation of 0.30 is regarded as a "medium" effect.[11] Therefore, we can conclude there is a medium empirical relationship between price and quality. More recently, Doucouliagos (2011) established guidelines for the practical significance of the partial correlation.[12] The partial correlation coefficient is considered "small" if the absolute value is between 0.07 and 0.17, and "large" if the absolute value is greater than 0.33. Therefore, according to these alternative criteria, a correlation of +0.30 might be regarded as approximately "large."

The existence of asymmetry in effect sizes is typical of meta-data sets, and from a visual examination of the funnel plot cannot be ruled out for our data. At least two possible explanations for asymmetries can be provided: *a*) publication selection bias; and/or *b*) heterogeneity across the employed data sets, and sample and modeling designs. We explore both possibilities below.

Meta-regression Analysis (MRA)

Our meta-regression analysis involves analyzing the distribution of the reported estimates

[8] The appendix indicates the magnitude of the multiple estimates emanating from the examined studies. In general, the multiple estimates do contain independent information, as they typically relate to either different time periods, variety/style categories, regions, and expert ratings. Effectively, the multiple estimates employ different information and hence multiple counting is not particularly evident in the meta-data set.

[9] Using the Fischer z-transformed partial correlations produces essentially the same funnel plot, except that there is a longer tail for some of the positive correlations that are estimated with lower precision.

[10] The weighted average is constructed by using the inverse variance of each partial correlation as the weight. The more precise estimates are assigned a higher weight in the overall average.

[11] According to Cohen's (1988) guidelines, a *simple* correlation of less than 0.1 is a small effect, while anything larger than 0.5 is a large effect.

[12] These guidelines are based on the distribution of over 22,000 reported partial correlations in empirical economics.

and identifying the factors that drive heterogeneity in this literature. We simultaneously explore whether the asymmetry in the correlation estimates (figure 1) can be attributed to publication bias. The MRA model is given by:

$$(1) \qquad r_i = \beta_0 + \Sigma\beta_k z_{ki} + \beta_1 se_i + \varepsilon_i$$

where r denotes the partial correlation between price and quality, z_k represent regressors that capture the variation in studies because of sample design and extraneous factors, se denotes the partial correlation's standard error, and ε denotes an error term. Equation (1) is known as the FAT-PET (Funnel Asymmetry Test – Precision Effect Test) MRA (Stanley 2005; 2008). Weighted least squares (WLS) (using $w = 1/se^2$) is employed to estimate equation (1) to capture differences in the error variance due to differences in the variances of partial correlation estimates across studies. It has been shown that the inverse variance results in "optimal weights" (Hedges and Olkin 1985).

The $\beta_1 se$ term captures the impact of publication selection bias. If a study is free of publication selection bias, then estimated effect sizes (here, partial correlations) will not be correlated with their standard errors (Egger et al. 1997; Stanley 2005; 2008). However, in the case of publication selection bias, researchers search for estimates that are statistically significant and re-estimate econometric models until the relationship between r and se achieves some "acceptable" standard of statistical significance (e.g., a statistical significance at the 10% level). Thus, this selection will produce a correlation between the partial correlations with their standard errors (Stanley 2008).

A test of $\beta_1 = 0$ (FAT) is a test of the existence of asymmetry in the estimates and publication selection. It is important to test for publication selection bias because statistical inferences may be erroneous if the reported estimates are a biased sample of all estimates. Much evidence exists to suggest that researchers prefer reporting certain results and suppressing others (Roberts and Stanley 2005; Stanley and Doucouliagos 2012). Typically, the bias is in favor of rejecting the null hypothesis of a zero effect. In such cases, authors do not report all the results they uncover. Rather, they select results that are consistent with their prior findings, or results that they believe have a stronger chance of being published. This process suppresses certain findings while others are over-represented. Thus, publication selection bias will tend to inflate the magnitude of an empirical effect.[13]

Empirical effects beyond publication selection in equation (1) are captured by ($\beta_0 + \Sigma\beta_k z_{ki}$). Testing the equivalence of the effect size to zero (PET) is the test of the existence of an underlying effect after correcting for any publication selection bias.

A potential estimation complication for equation (1) is the non-independence of error observations given that in some instances we have employed multiple estimates from the same study. This recognition necessitates the use of a cluster error estimator for WLS based on study stratification or other panel data estimators that account for within-study dependence. There is, as yet, no clear consensus on how clustering should be treated in very unbalanced data sets used in MRA. Some studies employ WLS with cluster-corrected standard errors, while others employ fixed effects (FE) or random effects (RE) panel data estimators (e.g., Stanley and Doucouliagos 2012).

We will employ various estimators for the WLS standard errors to cater for effects of error dependence because of study stratification. The standard robust, cluster robust, and bootstrap methods are used. In our study, the number of clusters is small (36 when study identification (ID) is used), which suggests the wild bootstrap may be preferable to other bootstrap methods (see Cameron, Gelbach, and Miller 2008).[14] Further, the size of the clusters varies significantly in our data set. Ten clusters have only a single observation, while five clusters have 10 or more observations. A recent study by MacKinnon and Webb (2014) suggests the wild bootstrap may perform well in cases where there is severe inequality between cluster sizes.

Many MRA studies have employed panel data estimators to account for study dependence, with the most commonly employed estimator being the RE model (Stanley and Doucouliagos 2013). The RE method uses modified inverse variance weights, $(1/se^2 + \tau^2)$, where τ^2 is the estimate of

[13] A referee points out that in the context of wine, publication bias in reporting significant results may be less pressing. Accepting the null of no correlation between wine price and quality produces "interesting and intriguing" results and possibly is consistent with the subjective nature of wine quality ratings.

[14] The operating rule of thumb is that 42 clusters are needed when using cluster robust standard errors, see Angrist and Pischke (2008).

random effects variance (the between-study or heterogeneity variance). However, there is criticism of this practice (Stanley and Doucouliagos 2012). For example, one problem with random effects is that the estimation depends critically on a reliable estimate of the random effects variance (τ^2), (see Hedges and Vevea 1998). Also, the RE specification assumes the random effects are independent of the MRA moderator variables (Nelson and Kennedy 2009). The violation of this assumption results in inconsistent estimates of the MRA regression coefficients.

The FE model overcomes some of the deficiencies of the RE model by explicitly permitting correlation between the study effects and moderator variables. However, the FE model has some limitations (Nelson and Kennedy 2009). For example, the FE model estimates an "effect parameter" for each cluster, and with many clusters this consumes a significant number of degrees of freedom. This has important implications for the efficiency of estimates.

When assessing the relative merits of WLS and the panel data estimators for MRA, Stanley and Doucouliagos (2013) show that WLS MRA is superior to both the FE and RE specifications. Compared to both FE and RE, WLS MRA provides more accurate estimates of meta-regression coefficients and confidence intervals when there is no publication bias, and is much more accurate when publication selection bias is present.[15]

Given the limitations of the RE and FE models, our preference is for the WLS estimates with clustered corrected standard errors. However, as recommended by Nelson and Kennedy (2009) and the MAER-NET guidelines (Stanley et al. 2013), to examine the robustness of the WLS estimates we will also estimate the RE and FE models using the meta-data set. We will also conduct extra robustness checks for the WLS estimates, including the robust regression, Fisher z correlations, and clustering by author ID.

We will consider a number of regressors (z_k in equation (1)) to recognize the differences between studies, which may explain the heterogeneity of partial correlation estimates. The summary statistics for the employed regressors are provided in table 1. We broadly group the attributes of the employed studies into structural and technical characteristics. The structural characteristics potentially have implications for consumers and producers, including attributes such as reputation (wine and producer), individual chemical and sensory attributes, grape variety/style, region, etc. Technical characteristics include the type of quality rating scale, functional form, and treatment of endogeneity.

Sample design issues are typically addressed by either limiting the sample to specific subsets of wines, for example, Bordeaux reds, or by examining a spectrum of wines and including variables such as variety, region, and vintage to control for wine differences. To the extent the use of control variables in a sample of a cross section of wines fails to adequately control for variations in the sampled wines, we consider several sample design dummy variables. The variable *Single Region* identifies studies that have used wines from a single region, for example, Napa Valley or Bordeaux, while *Single Variety/Style* identifies studies that examine a single variety or style, for example a focus on Cabernet Sauvignon or Champagne.

Sample design issues also extend to the coverage of vintages and the location of wines examined. There is insufficient detail in the available data to allow an extensive examination of vintage and location issues, and so two broad measures are employed. To capture the effect of the strength of the correlation over time, we employ *Vintage 2000*, which counts studies using data for vintages after the year 2000. To capture any difference that may exist because of the general location of produced wines, we employ *New World*, a variable which counts for studies that analyze wines not produced in Europe.

As previously indicated, many studies have argued that as wine is an experience good, consumers do not know its current quality, so the reputation of a wine better reflects the taste preferences of consumers. This argument potentially weakens the importance of sensory quality if the reputation of a wine is also included in the hedonic wine function (*Wine Reputation*). Typically, wine reputation is measured by using the average quality score of a number of previous vintages. The length of the lag varies from one period (Ramirez 2010) to all available past data (Ali and Nauges 2007). Given the expectation of a positive high correlation between sensory quality and wine reputation,

[15] The WLS MRA has lower mean square error and bias, and in most cases has better coverage probabilities.

Table 1. Quality Ratings and Wine Prices Meta-data Summary

Label	Definition	Mean	Standard Deviation
Effect Measure			
Partial Correlation	Partial correlation between quality ratings and wine prices	0.371	0.26
Standard Error	Standard error of the partial correlation estimates	0.061	0.05
Structural Characteristics			
Single Region	Model includes only wines from a single wine region = 1, otherwise = 0	0.288	0.45
Single Variety/Style	Model includes only wines that use a single variety/style = 1, otherwise = 0	0.326	0.47
Wine Reputation	Model includes a measure of the wine's reputation (based on the wine's quality from previous vintages) = 1, otherwise = 0	0.120	0.33
Producer Reputation	Model includes a measure of the wine producer's reputation (based on the range of quality wines previously produced) = 1, otherwise = 0	0.201	0.40
Sensory	Model includes measures of individual sensory variables (complexity, balance, etc) = 1, otherwise = 0	0.043	0.20
Chemical	Model includes measures of individual chemical variables (acid, sugar, etc) = 1, otherwise = 0	0.120	0.33
New World	Model includes only wines that are not produced in Europe = 1, otherwise = 0	0.565	0.50
Vintage 2000	Model includes wines for vintages after 2000 = 1, otherwise = 0	0.467	0.50
Technical Characteristics			
Endogenous	Model employs an estimation technique that recognizes endogeneity = 1, otherwise = 0	0.087	0.28
Rating Scales			
100 Scale*	Model employs a 100 point quality rating scale = 1, otherwise =0	0.728	0.45
20 Scale		0.076	
5 Scale		0.152	
Other		0.043	
Functional Form			
Log Linear		0.478	
Double Log*	Model employs the double-log functional form = 1, otherwise = 0	0.310	0.46
Linear Polynomial		0.087	
Other		0.125	
Prices			
RRP*	Model employs recommended retail prices = 1, otherwise = 0	0.804	0.40
Auction		0.109	
Average Retail		0.087	

Note: The N = 184 from 36 studies. Asterisk denotes the employed regressor.

then omitted variable bias arguments (e.g., Gujarati and Porter 2009) would suggest that both a smaller quality coefficient estimate and a larger quality standard error estimate may result from including wine reputation. This leads to a smaller price-quality partial correlation.[16]

Relatedly, we will consider *Producer Reputation* to capture the impact of the importance of the reputation of the producer in studies. We expect this variable to also weaken the price-quality relationship. Employed measures of producer reputation vary from the number of outstanding wines

[16] If wine reputation had little or no impact on price, then including an irrelevant variable occurs. Even though the coefficient estimate for quality may be unaffected, its standard error may still be higher.

produced by the vineyard in the past two years (Berrios and Saens 2012) to expert subjective ratings based on the producer's past quality ratings over three years (Schamel and Anderson 2003).

A minority (16%) of estimates have also included individual sensory (e.g., complexity, balance, and firmness) and/or chemical measurement (e.g., acid, sugar, and density) variables in price functions. Given the previously discussed, established relationships between expert wine scores and desirable sensory and chemical attributes, we expect the inclusion of these variables to weaken and make less precise the correlation between price and quality. To capture these effects, (*Sensory*) and (*Chemical*) variables are included in the MRA.

We now turn our attention to some of the technical characteristics of the examined studies. The bulk (73%) of the estimates employ a 100-point scale for wine assessment. Most other studies employ 20- or 5-point scales, (see Cicchetti and Cicchetti 2009) for a discussion of different wine rating scales. The gradation is much finer for the 100-point scale assessments and this may impact the estimated price-quality relationship. Assuming similar levels of wine assessment validity across scales, we expect that a scale with finer gradation may better capture the variation in quality, and hence price, than a scale with a coarser gradation. Taylor and Yu (2002) have shown how categorizing a continuous variable leads to lower levels of predictive ability. This may imply a loss of correlation strength in moving from a 100- to 5-point quality scale. The 100-point scale may better reflect the "true" underlying continuously measurable quality of a wine. To capture this possible effect, we use a dummy variable for studies that employ a 100-point wine quality scale (*100 Scale*).

Even though the majority (80%) of estimated hedonic models have employed RRPs, some have used either average retail or auction prices. Given that RRPs are typically set by wine producers as standard retail mark-ups over wholesale prices, then their exposure to current demand influences is expected to be weaker than average actual retail prices or auction prices. To this extent we expect that studies that use (*RRP*) may have a weaker price-quality correlation.

A small number (9%) of estimates have accounted for the possible endogeneity of the quality variable in price functions. Two arguments have been offered. First, some expert tasters may have some approximate knowledge of the price of the wine before assessing and scoring a wine. A two-stage least squares (2SLS) has been used to recognize this effect. Second, others have argued that the quality of a wine is a latent construct and can only be reflected by multiple observed indicators (scores from a number of guides). The 2SLS is also used here to explicitly recognize the consequent measurement error associated with observed indicators. Two possible opposing influences on the price-quality correlation from using 2SLS may exist. First, the use of ordinary least squares (OLS) in the presence of "error in variables" is known to have an attenuation effect on estimates. This biases the coefficient estimate toward zero (Wooldridge 2006). In other words, the 2SLS quality coefficient estimate may be higher as it accounts for the attenuation bias. On the other hand, 2SLS estimates are known to have higher standard error estimates than their OLS counterparts, particularly if the employed instruments are weakly correlated with the explanatory variables (Wooldridge 2006). These opposing effects may lead to any possibility for the difference in partial correlation estimates between 2SLS and OLS. In any event, we include the variable *Endogenous* to assess any potential impact. Finally, to account for any variability due to functional form, a variable is included for the use of a double-log function form (*Double Log*).

Results

Baseline OLS results and the WLS estimates using various standard error estimators for equation (1) are presented in table 2. The WLS standard errors are estimated using the following: robust column (2); cluster-robust by study (3); bootstrap (4); and wild bootstrap (5).[17]

The constant estimate in the MRA is 0.458, and suggests a large correlation.[18] The funnel

[17] The intraclass correlation is 0.573. This is fairly large and indicates that errors are correlated within clusters. Failing to account for this dependence can understate standard errors (Moulton 1990).

[18] For all models, the constant offers an estimate of the partial correlation: corrected for any selection bias; for wine that does not come from a single region; that is not a single variety; where ratings do not use the fine 100-point scoring; where wine and producer reputation is ignored; using the auction or average price

Table 2. Meta-regression Analysis of Quality Ratings and Wine Prices (Dependent Variable is Partial Correlations)

Variable	OLS (1)	WLS, Robust SE (2)	WLS, Cluster SE t-ratios (3)	WLS Bootstrap SE t-ratios (4)	WLS, Wild Bootstrap SE p-values (5)
Constant	**0.377**	**0.458**			
	(5.69)	**(6.90)**	**(5.76)**	**(2.21)**	**[0.000]**
Standard error	−0.237	−0.429			
	(−0.74)	(−0.82)	(−0.65)	(−0.46)	[0.505]
Single Region	−0.004	**–0.131**			
	(−0.10)	**(–2.48)**	**(–1.89)**	(−0.74)	[0.185]
Single Variety/Style	**0.230**	**0.161**			
	(5.86)	**(3.20)**	**(2.34)**	**(1.90)**	**[0.065]**
100 Scale	**0.098**	**0.127**			
	(3.08)	**(2.80)**	**(2.66)**	(1.52)	**[0.030]**
Wine Reputation	**–0.272**	**–0.220**			
	(–7.55)	**(–6.75)**	**(–4.38)**	**(–2.99)**	**[0.010]**
Producer Reputation	0.015	0.036			
	(0.44)	(1.19)	(0.99)	(0.44)	[0.565]
RRP	0.002	**–0.207**			
	(0.03)	**(–2.86)**	**(–2.52)**	(−0.99)	[0.105]
Sensory	**–0.124**	0.099			
	(–1.93)	(1.03)	(0.79)	(0.51)	[0.540]
Chemical	**0.190**	**0.237**			
	(3.48)	**(2.68)**	(1.56)	(0.98)	[0.390]
Endogenous	**0.085**	**0.176**			
	(2.00)	**(3.04)**	**(2.22)**	(1.41)	[0.345]
New World	**–0.112**	0.020			
	(–2.52)	(0.45)	(0.38)	(0.19)	[0.765]
Double Log	**–0.190**	**–0.122**			
	(–4.67)	**(–2.99)**	**(–2.37)**	(−1.17)	**[0.095]**
Vintage 2000	−0.025	−0.026			
	(−0.93)	(−0.67)	(−0.70)	(−0.41)	[0.570]
F-test	**32.95**	**22.09**	**17.15**	**111.26**	
Adjusted R^2	0.69	0.54	0.54	0.54	

Note: The N = 184, from 36 studies. Parentheses contain t-ratios and brackets present p-values. Column 1 reports OLS results without adjusting standard errors. Estimates reported in column 2 use WLS with the inverse variance used as weights. Columns 2 to 4 report t-ratios based on alternative standard error estimates: (2) robust; (3) adjusted for data clustering at the study level; (4) bootstrap. Column 5 presents the p-values for the WLS estimates using the Cameron, Gelbach, and Miller (2008) wild bootstrap. All bootstrapping uses 1,000 replications. Cell entries in bold denote statistical significance at least at the 10% level.

plot displays asymmetry that can be due to either selection bias or actual heterogeneity in the research process (recall the wide spread of results illustrated figure 1). The standard error variable is statistically insignificant for all WLS-estimated models in table 2. This implies there is no selection bias once the heterogeneity in study design is modeled. Hence, we conclude that the asymmetry in the funnel graph can be attributed to heterogeneity.

Four of the variables included in the MRA emerge as important explanatory variables of the observed heterogeneity: *Wine Reputation, Single Variety/Style, 100 Scale*, and *Double Log*. The WLS estimates for each of these variables are statistically significant when using either the robust, cluster, or wild bootstrap t-ratios. For the standard bootstrap t-ratios, only *Wine Reputation* and *Single Variety/Style* are statistically significant. As suggested previously, given the severe inequality between cluster sizes, the standard bootstrap t-ratios may underestimate the importance of variables compared to the wild bootstrap.

As expected, when a wine's reputation is included in a model, the partial correlation

of wine; ignoring individual sensory features; the wine is not from the new world; the vintage is pre-2000; and estimation does not use the double log form. Note, however, that this configuration applies to only a small subgroup of estimates.

between price and quality weakens; its point estimate of −0.220 is rather large in the context that the weighted average partial correlation for all models is 0.30. When a model focuses on a single variety or style, the price-quality correlation increases by 0.161. Using a 100-point scale for wine quality assessments increases the partial correlation by 0.127. The stronger partial correlations for both a single variety and 100-point scale possibly reflect a better capacity to measure quality, which strengthens its relationship with price. Finally, using the double-log functional form reduces the price-quality partial correlation by 0.122. This result possibly reflects the dominance of the better-fitting log-linear form in the majority of estimated models.

Three other variables have some moderate importance. *Endogenous, RRP*, and *Single Region* are statistically significant in two of the four WLS specifications and have marginal impacts over 0.1. Each of these variables is not identified as being significant when using the conservative bootstrap t-ratios, but are significant when employing more conventional robust or cluster t-ratios. For models that recognize endogeneity, partial correlations are estimated to be 0.176 lower. This implies the correction for possible attenuation bias increases the partial correlation. For models that employ recommended retail prices, partial correlations are estimated to be 0.207 lower. As expected, this suggests that actual retail and auction prices may better capture demand influences than RRPs. Actual retail and auction prices reflect better-informed quality assessments and hence their use results in stronger price-quality relationships. Finally, for models that employ wines from only one region, partial correlation estimates are 0.131 lower. This result may reflect the difficulty of comparing the qualities of different wine types from a single region. The result may also be an outcome of an inability of studies that use wines from multiple regions to control adequately for region effects by using dummy variables.

Using chemical variables in specifications has the largest estimated marginal impact (0.237). This variable is significant for the robust SE estimator, but it is not significant in any of the models which recognize the dependence of observations. The remaining four variables have marginal estimates of less than 0.1, and are never significant for the WLS estimates. This implies that partial correlations are not influenced by including producer reputation, individual sensory variables, wines from the New World, and more recent vintages.

Table 3 presents the robustness checks for the MRA results. Column (1) reports the results from the precision-effect estimate with standard error (PEESE) estimator, where the square of the standard error replaces the standard error. Stanley and Doucouliagos (2012) suggest that replacing *se* by se^2 will provide a more accurate estimate of the underlying effect and will better identify the publication selection bias in the studies. Column (2) reports the results of using robust regression, which corrects the estimates for the impact of any leverage points.[19] Column (3) uses the Fisher z-transformed correlations rather than the unadjusted correlations and the wild bootstrap for standard errors. Column (4) uses author ID rather than study ID to cluster the observations.

To further examine the robustness of the WLS estimates, alternative panel-data models are also estimated. Column (5) uses the random effects MRA estimated via restricted maximum likelihood (REML). Column (6) is a random effects panel data WLS model, while column (7) presents results from a fixed effects panel data WLS model. Both of these latter panel data models are weighted using inverse variance weights and the standard errors are adjusted for data clustering within studies.

Our preferred results are presented in table 2. In a broad sense the estimates in table 3 suggest that the results from table 2 are reasonably robust; however, there are some minor variations. The four variables identified as most important from the WLS models—*Wine Reputation, Single Variety/Style, 100 Scale*, and *Double Log*—in general appear to maintain their significance across the variety of alternative estimation procedures. Compared to the WLS estimates, *Single Region* and *RRP* appear to be more important for some of these alternative estimation procedures. These two variables are statistically significant in five of the alternative seven models. The remaining variables—*Producer Reputation, Sensory, Chemical, Endogenous, New World*, and *Vintage 2000*—continue to be relatively

[19] We actually estimate a WLS version by first dividing all variables by the standard error of the partial correlations, and then applying robust regression.

Table 3. Robustness Checks Meta-regression Analysis of Quality Ratings and Wine Prices (Dependent Variable is Partial Correlations)

Variable	WLS, PEESE Cluster SE (1)	WLS, Robust regression (2)	WLS, Fisher z Wild Bootstrap SE (3)	WLS, Wild Bootstrap SE author id (4)	REML (5)	Panel: Random effects, Cluster SE (6)	Panel: Fixed effects, WLS, Cluster SE (7)
Constant	**0.437**	**0.474**	**0.429**		**0.359**	**0.469**	**0.762**
	(5.97)	**(9.90)**	**[0.000]**	**[0.000]**	**(5.15)**	**(4.90)**	**(3.65)**
Standard error	–	−0.028	0.499		−0.123	−0.242	0.825
		(−0.07)	[0.600]	[0.490]	(−0.33)	(−0.23)	(0.45)
Single Region	**–0.125**	**–0.124**	**–0.131**		0.005	**–0.171**	**–0.594**
	(–1.81)	**(–3.45)**	**[0.235]**	[0.180]	(0.12)	**(–2.54)**	**(–3.89)**
Single Variety/Style	**0.148**	**0.234**	0.167		**0.228**	0.055	0.055
	(2.44)	**(7.53)**	[0.105]	**[0.035]**	**(5.95)**	(0.99)	(0.70)
100 Scale	**0.135**	**0.045**	**0.171**		**0.093**	**0.145**	**0.157**
	(3.05)	**(1.96)**	**[0.000]**	**[0.050]**	**(2.98)**	**(2.93)**	**(2.35)**
Wine Reputation	**–0.220**	**–0.170**	**–0.244**		**–0.270**	**–0.215**	**–0.297**
	(–4.38)	**(–10.05)**	**[0.015]**	**[0.005]**	**(–8.20)**	**(–3.32)**	**(–3.04)**
Producer Reputation	−0.033	**0.100**	0.030		0.012	0.017	0.043
	(0.91)	**(8.05)**	[0.605]	[0.530]	(0.37)	(0.33)	(0.39)
RRP	**–0.201**	**–0.206**	−0.203		0.031	**–0.250**	**–0.512**
	(–2.39)	**(–3.95)**	[0.180]	**[0.070]**	(0.41)	**(–3.28)**	**(–2.59)**
Sensory	0.092	−0.051	0.060		**–0.139**	0.084	**–0.523**
	(0.73)	(−0.83)	[0.770]	[0.550]	**(–2.10)**	(0.77)	**(–2.47)**
Chemical	0.223	**0.239**	0.380		**0.198**	0.211	**0.643**
	(1.41)	**(3.36)**	[0.480]	[0.425]	**(3.45)**	(0.94)	**(3.98)**
Endogenous	**0.173**	0.041	0.220		**0.091**	0.143	0.187
	(2.11)	(1.59)	[0.430]	[0.300]	**(2.31)**	(1.50)	(1.25)
New World	0.020	−0.034	0.008		**–0.121**	**0.075**	**0.055**
	(0.39)	(−1.35)	[0.885]	[0.860]	**(–2.82)**	**(2.52)**	**(2.46)**
Double Log	**–0.121**	**–0.089**	−0.146		**–0.208**	**–0.131**	**–0.175**
	(–2.35)	**(–3.59)**	[0.105]	[0.140]	**(–5.25)**	**(–2.70)**	**(–2.40)**
Vintage 2000	−0.024	−0.006	−0.025		−0.022	0.002	0.029
	(0.66)	(−0.44)	[0.620]	[0.545]	(−0.84)	(0.06)	(0.47)
Standard error Squared	−0.896	–		–	–	–	–
	(−0.22)						

Note: The N = 184 from 36 studies. Parentheses report t-ratios and brackets present p-values. Cell entries in bold denote statistical significance at least at the 10% level. Column 1 reports WLS estimates of the PEESE model. Column 2 reports results using Robust Regression. Column 3 uses the WLS estimates using the Fisher z-transformed partial correlations as the dependent variable using the wild bootstrap standard errors. Column 4 uses author IDs instead of study IDs to cluster-adjust the wild bootstrap standard errors using WLS. Column 5 reports results of the random effects of MRA estimated using restricted maximum likelihood, while column 6 reports the WLS random effects panel model. Column 7 is the WLS fixed-effect multilevel MRA. All WLS estimates use the inverse variance as weights.

unimportant. These variables are statistically significant in fewer than half of the alternative estimation procedures. The statistical insignificance of the standard error and standard error squared variables confirms the absence of any publication selection bias across these alternative estimation procedures.

In general, the panel-data models produce more statistically significant estimates than the WLS models. The fixed effects panel data representation is most different from the WLS estimates. The FE model suggests that eight of the variables are statistically important, and it identifies many unrealistically large marginal impacts for partial correlations. This occurs because the fixed-effects panel data model is capturing the within-study heterogeneity. For our data, the Hausman test suggests preferring random effects to fixed effects ($\chi^2 = 18.02$, p-value = 0.16) so we can reject the fixed effects panel estimates. Also, ten studies in our data set have single partial correlation estimates. The reliability and plausibility of a fixed effects specification with so many "single observation" fixed effect estimates is questionable (Nelson and Kennedy 2009).

These concerns cast some doubt about the improved importance of the *Single Region* and *RRP* variables in table 3, and reaffirm our preference for the WLS estimates in table 2.

Discussion and Implications

The main findings of the research indicate a moderate partial correlation between wine prices and sensory quality ratings; the precision weighed effect size is +0.30. The MRA identifies no publication selection bias across studies and attributes the asymmetry in partial correlation estimates to heterogeneity across a series of study design and control variables.

The moderate size of the partial correlation suggests some interesting implications. In the general commodity context, Caves and Greene (1996) suggest that a perfect correlation between price and quality mimics the full information case where full knowledge of the quality of the product is available to both consumers and producers. Deviations from a perfect correlation suggest the existence of imperfect information about quality and the scope for various strategies for both consumers and producers.

To the extent that a moderate rather than a zero partial correlation has been identified across studies implies that elements of full information strategies may be potentially important for wine consumers and producers. As suggested by Arias-Bolzmann et al. (2003), a high price-quality correlation implies that producers may be able to gain high rewards through vertical differentiation by improving quality that is under their control. A high degree of correlation also implies that price signals quality and consumers can purchase wine based on price with some certainty knowing that it is associated with an appropriate quality.

The extent to which the observed correlation deviates from unity also presents strategic possibilities for both consumers and producers. Profitable activities may emerge that take advantage of the lack of information about wine quality. Better-informed buyers could potentially identify bargains in the short run (Miller, Genc, and Driscoll 2007). In this context expert wine guides potentially play an important role (Oczkowski 2010). Indeed, in analyzing the lack of consensus between expert ratings, Cicchetti and Cicchetti (2013) note the comments of the wine expert Jancis Robinson, who suggests that individual consumers may wish to follow the "preferences and prejudices" of a specific wine critic in making wine purchase choices. It could be argued that this is an implication of preference heterogeneity and horizontal differentiation practices.

For producers, a whole range of strategies and implications emerge. Daughety and Reinganum (2008) and Schnabel and Storchmann (2010) provide a summary of the strategic producer literature that recognizes the existence of varying product quality in the presence of incomplete information and various market structures. For example, some low quality producers may be able to charge higher prices than implied by quality, in the short term, as buyers may find it uneconomic to conduct the necessary search to identify quality (Caves and Greene 1996). Nelson (1970) suggests that high-quality producers may set short-term prices below full-information levels (forgoing short-term profits) in an effort to scare off low-quality producers in the long run. Bagwell and Riordan (1991) develop a model that allows high-quality producers to charge above full-information price levels given that low-quality producers cannot sustainably follow a similar strategy. In summary, the specific strategies to follow will depend on the quality of the wines produced and the market structure.

Our research has also identified the importance of some key structural variables in implementing and designing wine price-quality studies. The most dominant variable in the MRA appears to be the inclusion of wine reputation. In nine of the 22 (41%) models that included both sensory quality and wine reputation variables, quality was found to be statistically insignificant and practically far less important than wine reputation (see Landon and Smith 1997; Oczkowski 2001; and Benfratello, Piacenza, and Sacchetto 2009). The relative unimportance of quality ratings in the presence of wine reputation variables possibly reflects the relative (to supply) importance of demand influences on prices. It is often argued that as an experience good, consumers lack information about the quality of a wine. As a result, consumers rely on the wine's reputation, as reflected by the quality of previous vintages, to make purchase decisions. This observation does not necessarily make information about the quality of a wine redundant when used with wine reputation information. Typically, quality ratings and wine reputation variables are highly correlated, but when they are not for a particular wine, potential bargains for consumers may be identified, that is, wines that have low reputation scores but high (current) quality ratings. Benfratello, Piacenza, and Sacchetto (2009) suggest that producers should aim at building an established reputation for the product. Potentially, this can be supported by pursuing promotional activities that may lead to wine's inclusion in expert guides. Apart from using promotional activities to raise a wine's profile, the employed wine reputation measures in the studies make it clear that it is the sustained sensory quality of a wine over time that leads to its high reputation. Wine producers could direct resources to activities that preserve this lasting quality.

It appears that the strength of the price-quality relationship is unaffected by including variables measuring the reputation of producers across studies. This does not imply that producer reputation is unimportant in explaining prices. In fact, in the 37 models that included producer reputation, the variable is statistically significant in 81% of cases. This implies that the impact of producer reputation does not necessarily occur through the price-quality relationship. Rather, its impact occurs directly on price. This suggests that strategies directed at improving producer reputation through producing a range of high-quality wines should not necessarily cease.

The relative unimportance of individual chemical and sensory measures is affirmed. This suggests that no added (beyond the overall sensory quality rating) impact on price exists with these variables. However, we did identify a relatively large (but imprecise) positive point estimate for the inclusion of chemical attributes. This may indicate some potential additional information contained in chemical compared to sensory measurements for quality. This needs to be confirmed through more appropriately-performed studies. We found no important difference in the strength of the price-quality relationship between new and old world wines across studies. This result does not necessarily contradict the importance of regionality for old world wines in influencing prices (Schamel 2006; 2009). The importance of regionality for old world wines appears to occur independently of the price-quality relationship.

Turning to some technical characteristics of the studies, the importance of both the use of the 100-point quality rating scale and studies that focus on single varieties/styles probably reflects an improved ability to measure the quality of a wine. An improved measure of quality enhances the employed variables' ability to be reflected in price variations, as it better captures demand and supply influences. The finer gradation of the 100-point scale compared to those that use 20 or 5 points possibly better reflects the underlying continuous quality variable. A focus on a single variety/style mimics the typical show judging system where comparisons are made between wines of the same variety/style in awarding medals and prizes. For example, Ashton (2013) presents evidence of a higher consensus measure between experts when focusing only on Bordeaux reds. In other words, it appears easier to determine the quality of wine within a variety rather than across varieties. This better quality assessment may enhance price variations. An alternative explanation for the estimated stronger partial correlation for a single variety/style relates to the horizontal differentiation strategies of producers. Producers may account for the importance of the price-quality relationship for particular wine varieties or styles when developing their product range.

The most interesting aspects of the other results suggest that stronger partial correlations are associated with the predominant log-linear form in modeling. It appears that functional form does matter. Some moderate evidence suggests that auction and average actual retail prices may better reflect demand influences on prices than the use of recommended retail prices. Also, some mild downward attenuation bias may not recognize the endogeneity of the quality ratings. The findings regarding the use of actual prices and endogeneity are not robust given the small number of studies that employ these features. Any definitive statements need to be confirmed by more research.

In conclusion, the systematic examination of over 180 hedonic wine price models developed over 20 years indicates the existence of a moderate and statistically significant partial correlation between wine prices and sensory quality ratings. Our finding suggests that when overall wine sensory ratings are employed, quality does matter to some extent. Quality matters, even given the subjective and inconsistent nature of expert ratings. The meta-regression analysis can explain over 50% of the heterogeneity in the estimated correlations by focusing on some key research design factors. Identifying these factors should be carefully considered by researchers when designing and developing hedonic wine price models.

References

Ali, H.H., and C. Nauges. 2007. The Pricing of Experience Goods: The Example of En Primeur Wine. *American Journal of Agricultural Economics* 89 (1): 91–103.

Angrist, J.D., and J-S. Pischke. 2008. *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton: Princeton University Press.

Arias-Bolzmann, L., O. Sak, A. Musalem, L. Lodish, K.R. Baez, and L.J. De Sousa. 2003. Wine Pricing: The influence of Country of Origin, Variety, and Wine Magazine Ratings. *International Journal of Wine Business Research* 15 (2): 47–57.

Ashenfelter, O. 2008. Predicting the Quality and Prices of Bordeaux Wine. *Economic Journal* 118 (529): F174–F184.

Ashton, R.H. 2012. Reliability and Consensus of Experienced Wine Judges: Expertise Within and Between? *Journal of Wine Economics* 7 (1): 70–87.

Ashton, R.H. 2013. Is there Consensus Among Wine Quality Ratings of Prominent Critics? An Empirical Analysis of Red Bordeaux, 2004–2010. *Journal of Wine Economics* 8 (2): 225–234.

Bagwell, K., and M.H. Riordan. 1991. High and Declining Prices Signal Product Quality. *American Economic Review* 81 (1): 224–239.

Ballester, J., B., Patris, R. Symoneaux, and D. Valentin. 2008. Conceptual vs. Perceptual Wine Spaces: Does Expertise Matter? *Food Quality and Preference* 19 (3): 267–276.

Benfratello, L., M. Piacenza, and S. Sacchetto. 2009. Taste or Reputation: What Drives Market Prices in the Wine Industry? Estimation of a Hedonic Model for Italian Premium Wines. *Applied Economics* 41 (17): 2197–2209.

Berrios, R., and R. Saens. 2012. The Country Brand Trap. *CEPAL Review* 106: 75–88.

Cameron, A.C., J.B. Gelbach, and D.L. Miller. 2008. Bootstrap-based Improvements for Inference with Clustered Errors. *Review of Economics and Statistics* 90 (3): 414–427.

Caves, R.E., and D.P. Greene. 1996. Brands' Quality Levels, Prices, and Advertising Outlays: Empirical Evidence on Signals and Information Costs. *International Journal of Industrial Organization* 14 (1): 29–52.

Charters, S., and S. Pettigrew. 2007. The Dimensions of Wine Quality. *Food Quality and Preference* 18 (7): 997–1007.

Cicchetti, D.V., and A.F. Cicchetti. 2009. Wine Rating Scales: Assessing their Utility for Producers, Consumers and Oenologic Researchers. *International Journal of Wine Research* 1: 73–83.

Cicchetti, D.V., and A.F. Cicchetti. 2013. As Wine Experts Disagree, Consumers' Taste Buds Flourish: How Two Experts Rate the 2004 Bordeaux Vintage. *Journal of Wine Research* 24 (4): 311–317.

Cohen, J. 1988. *Statistical Power Analysis in the Behavioral Sciences, 2nd edition*. Hillsdale, NJ: Erlbaum.

Combris, P., S. Lecocq, and M. Visser. 1997. Estimation for a Hedonic Price Equation for Bordeaux Wine: Does Quality Matter? *Economic Journal* 107 (441): 390–402.

D'Alessandro, S., and A. Pecotich. 2013. Evaluation of Wine by Experts and Novice Consumers in the Presence of Variations in Quality, Brand and Country of Origin Cues. *Food Quality and Preference* 28 (1): 287–303.

Daughety, A.F., and J.F. Reinganum. 2008. Imperfect Competition and Quality Signaling. *RAND Journal of Economics* 39 (1): 163–183.

Diaz-Plaza, E.M., J.R. Reyero, F. Pardo, G.L. Alonso, and M.R. Salinas. 2002. Influence of Oak Wood on the Aromatic Composition and Quality of Wines with Different Tannin Contents. *Journal of Agricultural and Food Chemistry* 50 (9): 2622–2626.

Djankov, S., and P. Murrell. 2002. Enterprise Restructuring in Transition: A Quantitative Survey. *Journal of Economic Literature* 40 (3): 736–792.

Doucouliagos, H. 1995. Worker Participation and Productivity in Labor-managed and Participatory Capitalist Firms: A Meta-analysis. *Industrial and Labor Relations Review* 49 (1): 58–77.

Doucouliagos, H. 2011. How Large is Large? Preliminary and Relative Guidelines for Interpreting Partial Correlations in Economics. *Economics Series Working Paper* 05, Deakin University.

Doucouliagos, H., and T.D. Stanley. 2013. Are All Economic Facts Greatly Exaggerated? Theory, Competition and Selectivity. *Journal of Economic Surveys* 27 (2): 3169–3339.

Edwards, F., and C. Mort. 1991. The Expert Wine Taster. *International Marketing Review* 8 (4): 8–12.

Egger, M., G.D. Smith, M. Schneider, and C. Minder. 1997. Bias in Meta-analysis Detected by a Simple, Graphical Test. *British Medical Journal* 315 (7109): 629–634.

Estrella Orrego, M.J.E., E. Defrancesco, and A. Gennari. 2012. The Wine Hedonic Price Models in the "Old and New World": State of the Art. *Journal of the Faculty of Agricultural Sciences, UNCuyo* 44 (1): 205–220.

Fogarty, J. 2003. Australian Premium Wine Prices: Do Quality Ratings Matter? *Western Australian Quarterly Bulletin of Economic Trends* 4 (3): 41–64.

Friberg, R., and E. Gronqvist. 2012 Do Expert Reviews Affect Demand for Wine? *American Economic Journal – Applied Economics* 4 (1): 193–211.

Garde-Cerdan, T., and C. Ancin-Azpilicueta. 2006. Review of Quality Factors on Wine Ageing in Oak Barrels. *Trends in Food Science and Technology* 17 (8): 438–447.

Gawel, R., A. Royal, and P.A. Leske. 2002. The Effect of Different Oak Types of the Sensory Properties of a Chardonnay Wine. *Australian and New Zealand Wine Industry Journal* 17 (1): 14–20.

Goldstein, R., J. Almenberg, A, Dreber, J.W. Emerson, A. Herschkowitsch, and J. Katz. 2008. Do More Expensive Wines Taste Better? Evidence from a Large Sample of Blind Tastings. *Journal of Wine Economics* 3 (1): 1–9.

Gujarati, D.N., and D.C. Porter. 2009. *Basic Econometrics, 5th edition*. New York: McGraw-Hill.

Hedges, L.V., and I. Olkin. 1985. *Statistical Methods for Meta-analysis*. Orlando, FL: Academic Press.

Hedges, L.V., and J.L. Vevea. 1998. Fixed- and Random-effects Models in Meta-analysis. *Psychological Methods* 3 (4): 486–504.

Hodgson, R.T. 2008. An Examination of Judge Reliability at a Major U.S. Wine Competition. *Journal of Wine Economics* 3 (2): 105–113.

Hunter, J.E., and F.L. Schmidt. 2004. *Methods of Meta-analysis: Correcting Error and Bias in Research Findings*. New York: Sage.

Jackson, D.I., and P.B. Lombard. 1993. Environmental and Management Practices Affecting Grape Composition and Wine Quality—A Review. *American Journal of Enology and Viticulture* 44 (4): 409–430.

Johnson, T., and J. Bruwer. 2004. Generic Consumer Risk-reduction Strategies (RRS) in Wine-related Lifestyle Segments of the Australian Wine Market. *International Journal of Wine Marketing* 16 (1): 5–35.

Landon, S., and C.E. Smith. 1997. The Use of Quality and Reputation Indicators by Consumers: The Case of Bordeaux Wine. *Journal of Consumer Policy* 20 (3): 289–323.

———. 1998. Quality Expectations, Reputation, and Price. *Southern Economic Journal* 64 (3): 628–647.

Lattey, K.A., B.R. Bramley, and T.L. Francis. 2010. Consumer Acceptability, Sensory Properties and Expert Quality Judgment of Australian Cabernet Sauvignon and Shiraz Wines. *Australian Journal of Grape and Wine Research* 16 (1): 189–202.

Lecocq, S., and M. Visser. 2006. What Determines Wine Prices: Objective vs. Sensory Characteristics. *Journal of Wine Economics* 1 (1): 42–56.

MacKinnon, J.G., and M.D. Webb. 2014 Wild Bootstrap Inference for Wildly Different Cluster Sizes. Queen's Economics Department Working Paper, No. 1314.

Mastrobuoni, G., Peracchi, F., and A. Tetenov. 2013. Price as a Signal of Product Quality: Some Experimental Evidence. *Carlo Alberto Notebooks, Working Papers.*

McConnell, K.E., and I.E. Strand. 2000. Hedonic Prices for Fish: Tuna Prices in Hawaii. *American Journal of Agricultural Economics* 82 (1): 133–144.

Miller, J.R., I. Genc, and A. Driscoll. 2007. Wine Price and Quality: In Search of a

Signaling Equilibrium in 2001 California Cabernet Sauvignon. *Journal of Wine Research* 18 (1): 35–46.
Moulton, B.R. 1990. An Illustration of a Pitfall in Estimating the Effects of Aggregate Variables on Micro Unit. *The Review of Economics and Statistics* 72 (2): 334–338.
Nelson, J.P., and P.E. Kennedy. 2009. The Use (and Abuse) of Meta-analysis in Environmental and Natural Resource Economics: An Assessment. *Environmental and Resource Economics* 42 (3): 345–377.
Nelson, P. 1970. Information and Consumer Behaviour. *Journal of Political Economy* 78 (2): 311–329.
Oczkowski, E. 1994. A Hedonic Wine Price Function for Australian Premium Table Wine. *Australian Journal of Agricultural Economics* 38 (1): 93–110.
Oczkowski, E. 2001. Hedonic Wine Price Functions and Measurement Error. *Economic Record* 77 (239): 374–382.
Oczkowski, E. 2010. Hedonic Wine Price Predictions and Non-normal Errors. *Agribusiness* 26 (4): 519–535.
Parr, W.V., K.G. White, and D.A. Heatherbell. 2004. Exploring the Nature of Wine Expertise: What Underlies Wine Experts' Olfactory Recognition Memory Advantage? *Food Quality and Preference* 15 (5): 411–426.
Ramirez, C.D. 2010. Do Tasting Notes Add Values? Evidence from Napa Wines. *Journal of Wine Economics* 5 (1): 143–163.
Reynolds, A.G., ed. 2010. *Managing Wine Quality, volumes I & II*. Oxford: Woodhead Publishing.
Roberts, C.J., and T.D. Stanley. 2005. *Meta-regression Analysis*. Oxford: Blackwell.
Rosen, S. 1974. Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition. *Journal of Political Economy* 82 (1): 34–55.
Saenz-Navajas, M-P., Y-S. Tao, M. Dizy, V. Ferreira, and P. Fernandez-Zurbano. 2010. Relationship Between Nonvolatile Composition and Sensory Properties of Premium Spanish Red Wines and their Correlation to Quality Perception. *Journal of Agricultural and Food Chemistry* 58 (23): 12407–12416.
Schamel, G., and K. Anderson. 2003. Wine Quality and Varietal, Regional and Winery Reputations: Hedonic Prices for Australia and New Zealand. *Economic Record* 79 (246): 357–369.
Schamel, G. 2006. Geography versus Brands in a Global Wine Market. *Agribusiness* 22 (3): 363–374.
Schamel, G. 2009. Dynamic Analysis of Brand and Regional Reputation: The Case of Wine. *Journal of Wine Economics* 4 (1): 62–80.
Schiefer, J., and C. Fischer. 2008. The Gap Between Wine Expert Ratings and Consumer Preferences: Measures, Determinants and Marketing Implications. *International Journal of Wine Business Research* 20 (4): 335–351.
Schnabel, H., and K. Storchmann. 2010. Prices as Quality Signals: Evidence from the Wine Market. *Journal of Agricultural & Food Industrial Organization* 8, 2.
Smart, R.E. 1985. Principles of Grapevine Canopy Microclimate Manipulation with Implications for Yield and Quality. A Review. *American Journal of Enology and Viticulture* 36 (3): 230–239.
Stanley, T.D. 2005. Beyond Publication Bias. *Journal of Economic Surveys* 19 (3): 309–345.
Stanley, T.D. 2008. Meta-regression Methods for Detecting and Estimating Empirical Effects in the Presence of Publication Selection. *Oxford Bulletin of Economics and Statistics* 70 (1): 103–127.
Stanley, T.D., and H. Doucouliagos. 2010. Picture This: A Simple Graph that Reveals Much Ado About Research. *Journal of Economic Surveys* 24 (1): 170–191.
Stanley, T.D., and H. Doucouliagos. 2012. *Meta-regression Analysis in Economics and Business*. Routledge: Oxford.
Stanley, T.D., and C. (H). Doucouliagos. 2013. Better than Random: Weighted Least Squares Meta-regression Analysis. *School Working Papers, Economics Series 2013-2*, Deakin University.
Stanley, T.D., H. Doucouliagos, M. Giles, J. Heckemeyer, R.J. Johnston, P. Laroche, J.P. Nelson, M. Paldam, et al. 2013. Meta-analysis of Economics Research Reporting Guidelines. *Journal of Economic Surveys* 27 (2): 390–394.
Taylor, J.M.G., and M. Yu. 2002. Bias and Efficiency Loss due to Categorizing an Explanatory Variable. *Journal of Multivariate Analysis* 83 (1): 248–263.
Teuber, R., and R. Herrman. 2012. Towards a Differentiated Modeling of Origin

Effects in Hedonic Analysis: An Application to Auction Prices of Specialty Coffee. *Food Policy* 37 (6): 732–740.
Varela, P., and A. Gambaro. 2006. Sensory Descriptive Analysis of Uruguayan Tannat Wine: Correlation to Quality Assessment. *Journal of Sensory Studies* 21 (2): 203–217.
Verdu Jover, A.J., F.J. Llorens Montes, and M.M. Fuentes Fuentes 2004. Measuring Perceptions of Quality in Food Products: The Case of Red Wine. *Food Quality and Preference* 15 (5): 453–469.
Volckner, F., and J. Hofmann. 2007. The Price-perceived Quality Relationship: A Meta Analytic Review and Assessment of its Determinants. *Marketing Letters* 18 (3): 181–196.
Wade, C. 1999. Reputation and its Effect on the Price of Australian Wine. *Australian and New Zealand Wine Industry Journal* 14 (4): 82–84.
Waugh, F.V. 1928. Quality Factors Influencing Vegetable Prices. *American Journal of Agricultural Economics* 10 (2): 185–196.
Weekly Times. 2013. *Price Rise for Penfolds Grange*. April 4.
Wooldridge, J.M. 2006. *Introductory Econometrics: A Modern Approach, 3rd edition*. Mason OH: Thomson South Western.
Zhao, W. 2008. Social Categories, Classification Systems, and Determinants of Wine price in the California and French Wine Industries. *Sociological Perspectives* 51 (1): 163–199.